Assignment 3

HUZAIFA JAHAN (215668106)
huzaifaj@yorku.ca

March 20, 2021

Note: This assignment is mainly for you to review several important discriminative models. You have to work individually. **You must use the same mathematical notations in textbook or lecture slides to answer these questions.** You must use this latex template to write up your solutions. Remember to fill in your information (name, student number, email) at above. No handwriting is accepted. Direct your queries to Hui Jiang (hj@eecs.yorku.ca)

Exercise 1

(30 marks) **Fully-Connected Neural Networks**

(a) (5 marks) Q8.2 on page 199 (see the margin note on page 175 for some examples.)

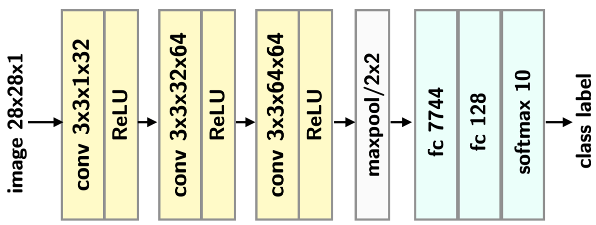
(b) (25 marks) You will use the MNIST data set for this question. Implement the forward and backward passes for fully-connected deep neural networks as in Figure 8.19. Use all MNIST training data to learn a 10-class classifier using your own back-propagation implementation, investigate various network structures (such as different number of layers and nodes per layer), and report the best possible classification performance in the held-out MNIST test images. Note that you are only allowed to use libraries for linear algebra operations, such as matrix multiplication, matrix inversion, and etc. You are not allowed to use any existing machine learning or statistics toolkits or libraries or any open-source code for this question.

Exercise 2

(30 marks) **Convolutional Neural Networks**
Note that 4404 students are required to do part (b) only (30 marks) while 5327 students need to do both parts (a) and (b) (10 marks + 20 marks).

(a) Q8.6 on page 200

(b) Implement the forward and backward passes for the following convolutional neural network:



Use all MNIST training data to learn a 10-class classifier and report the best possible classification performance in the held-out MNIST test images. Note that you are allowed to use any machine learning or statistics toolkits or libraries for this question. Do some investigations to ensure you use a suitable toolkit for this question.

Exercise 3

(20 marks) **Ensemble Learning**

(a) (10 marks) Adaboost: Q9.1 on page 215

$$\varepsilon = 0.3$$
$$\alpha = 1/2^{\star} \ln(1 - 0.3/0.3) = 0.42$$
$$\varepsilon = 0.7$$
$$\alpha = 1/2^{\star} \ln(1 - 0.7/0.7) = -0.42$$
$$\varepsilon = 0.5$$
$$\alpha = 1/2^{\star} \ln(1 - 0.5/0.5) = 0$$

In this example we used the

$$f_m(\mathbf{x})$$

classifier resulting in this example error set. Accuracy higher than 50% results in a positive weight for the classifier. Classifier's with exactly 50% accuracy is 0 this would hold for the

$$\bar{f}_m(\mathbf{x}) = -f_m(\mathbf{x})$$

as it would be the same point , and thus, does not contribute to the final prediction. Finally errors 0.3 and 0.7 lead to classifier weights with inverse signs which would lead to equivalent use of either

$$f_m(\mathbf{x})$$
$$-f_m(\mathbf{x})$$

(b) (10 marks) Gradient Tree Boosting: Q9.5 on page 215

What to submit?

You must submit:

1. one PDF document (using this latex template) for your solutions to all written questions and all results and discussions for your programming assignments;
2. one zip file that includes all of your Python codes (e.g., *.ipynb if you use Jupyter notebooks) and a readme file for TA to run your codes;

from eClass before the deadline. No late submission will be accepted.